

# First Survey Effort on Translation Systems

Pramod Salunkhe<sup>1</sup>, Aniket Kadam<sup>2</sup>, <sup>3</sup>Rutuja patil

<sup>1</sup>Savitribai Phule Pune University, Dept of Computer Engineering  
Lavale pune India

<sup>2</sup>Bharati Vidyapeeth University, Dept of Information Technology,  
Pune, Katraj

<sup>3</sup>Bharati Vidyapeeth University, Dept of Information Technology,  
Pune, Katraj

## Abstract

*Translation is required to eliminate communication difference that occurs in between two languages and discard them from sharing information in between them.. Translation is been require by authors for writing rich literature work in one language to other and hence need a faster and better translators. Translation is done with human translators and need to be paid heavily and is time consuming process. Translation is need at airport when priministers from different centuries meet. At travel destination where we are new. Computer help in summarizing huge calculations and complex problems in seconds, which are been extended to solve this problem and called as machine assisted or machine translation. Huge scope of work lies in development of these machines. First step in better research identification is systematic survey. The paper focuses on survey of translation systems currently available merits demerits and scope of work to be done. This article is systematic survey on 10 articles and first survey work on “Marathi Translation systems”.*

**Keywords:** Machine Translation, Marathi, Hindi, Google translator, Natural language processing, RBT (rule based Translation), SBMT (statistical based Translation), HMT (Hybrid machine translation).

## 1. INTRODUCTION

Machine translation eradicates gap in communication and facilities sharing information among people [1]. Sanskrit has been the mother language and Marathi Tamil are siblings of Sanskrit. Many other more than 70 million languages have been derived from all above three[1]. Major problem in developing translators is huge vocabulary and structural divergence introduced by language. The main goal of research is to focus on English and Marathi language as English is worldwide accepted and Marathi is our mother tongue India is a multilingual country where the spoken language changes after every 50 miles. There are 22 official languages; and approximately 2000 dialects are spoken by different communities in India[5]. English and Hindi are used for

official work in most states of India.

The state governments in India predominantly carry out their official work in their respective regional language whereas the official work of Union government is carried out in English and/or Hindi. All the official documents and reports of Union government are published in English or Hindi or in both English and Hindi. Many newspapers are also published in regional languages.[3,4] Translating these documents manually is very time consuming and costly. Hence there is need to develop good machine translation (henceforth referred as MT) systems to address all these issues, in order to establish a better communication between states and Union governments and exchange of information amongst the people of different states with different regional languages.

Indian languages are divided into five language families; viz. Indo-Aryan (76.87% speakers), Dravidian (20.82% speakers), Austro-Asiatic (1.11% speakers), Tibeto-Burman (1% speakers) and Andmanese (less than 0.001% speakers). Many Indian languages being low resource languages become a major hurdle in the development of MT systems for Indian languages.[8,9]

Many researchers, institutions and research organizations in India have started working on MT systems for English to Indian languages and among Indian languages have succeeded in obtaining very satisfactory results.[9,10].

**The paper is been organized in 4 section. Section I on Introduction II on Literature Survey III on issues and challenges IV on Conclusion and scope.**

**2. LITERATURE SURVEY**

Author	Abstract and Methodology	Scope
<b>Abhay</b>	Assertive sentence translation from English to Marathi. Tokenization, POS. Rule set for English to Marathi, with Bilingual Dictionary is technique used in system	Only Assertive sentences have been taken can be worked for all 5 types of sentences Interrogative etc.
<b>Charugatra</b>	Proposed work is on inflection rules which highly give fluency to sentence or speech we speak. Core technique used for this is SOV and SVO structural Writing for rule generation.	Writing rule is time and costly work we need better Technique to solve this problem.
<b>Dhore</b>	<p>Focuses on Difficulties faced due to language and presents issues and challenges in common</p> <ul style="list-style-type: none"> <li>❖ Script Specification</li> <li>❖ Missing Sounds</li> <li>❖ Multiple Transliterations</li> <li>❖ Spelling by orthography or Phonology</li> <li>❖ Allophones</li> <li>❖ spelling Variations</li> <li>❖ Capitalization</li> <li>❖ Language of Origin</li> <li>❖ Short Vowel and Long Vowel Conjuncts</li> <li>❖ Affixes</li> <li>❖ Acronyms</li> <li>❖ Loan Words</li> <li>❖ Incorrect Syllabification</li> <li>❖ Consonant 'r' in the conjuncts</li> <li>❖ Schwa Identification and Deletion</li> </ul>	Transliteration is first step for better translation. Using tag we can reduce errors in system.
<b>Rajat</b>	Work is on Sysnet for three works on Cross lingual IR, English to Indian MT, Indian languages to Indian languages MT. proposed system uses SYSNET not just words but concept behind them. Core technique is concept to concept mapping.	Development of wordnet would hence target lab]language translation like Marathi
<b>Nikhil</b>	Mobile are been used in daily life even for small task. Translation	The System can be made better with more techniques from rich

	of complete sentence bring error. proposed work presents using common words to present user need. And use simplified approach.	IR domain techniques like feedback mechanism etc.
<b>Jyoti</b>	Rich language like Marathi requires POS .proposed system is statistical trigram POS Tagger, exploring most likely token is base of POS development. Trigram[w1,w2,w3]	Training system with mores examples would increase accuracy.
<b>Mangala</b>	CLIR MLIR various active area of Information retrieval in context to Marathi have been presented. MT System is base to achieve multi-lingual or cross lingual IR.	Good survey but lacks evaluation based on parameters and mathematical values.
<b>Sreelekha</b>	Comparative examination on statistical and rule based system . RMBT are accurate but require time and cost whereas SMBT is better in terms of fluency and lacks accuracy. Hence HMT would be best approach is system development.	Hybrid Translation is future scope .
<b>Pramod</b>	There System for rule based is been demonstrated with mapping methodology to map rule from English to Marathi. Even though methodology is effective it time consuming and limited	Rule based are limited systems and hence author gives Hybrid as future work scope.
<b>Pramod</b>	Hybrid system has been developed by author and performance is evaluated where performance of hybrid is found to be best when RBMT and SMBT is combined.	Though the system has limitations due to Domain RMBT and corpus of SMBT author highlight work of Pushpak Bhattacharya professor from IIT and concludes WORDNET for Marathi language as scope.

### 3 ISSUES CHALLENGES AND SCOPE

#### 3.1 Issues and Challenges

Several issues arise with different language and them as

- Word order structure.
- Parts of speech related to language
- SVO and SOV structure of Languages like English and Marathi, Hindi.
- RMBT is limited due to Rule writing and SMBT lacks

accuracy due to word to phrase replacement

#### 3.2 Scope

HMT (Hybrid machine Translation) is scope of work as its also has challenges and need to address areas like complete Trainable system and Implementation and active research area like wordnet are scope.[8,9,10]

#### **4 CONCLUSION AND FUTURE WORK**

The Current work in Marathi Translation is definitely Hybrid systems and development of WorldNet and POS taggers for rich language like Marathi This is survey 1 on Marathi system further we would like to extend it to survey 2 article for more detail examination.

#### **Acknowledgement**

I acknowledge every author whose article has been taken in citation and Prof Phushpak Bhattacharya for his work in MT prof. Shashank Joshi for guidance and all my friends

#### **References**

- [1] **Charugatra Tidke**, Shital Binayakya, Shivani Patil Rekha Sugandhi , “Inflection Rules for English to Marathi Translation”, International Journal of Computer Science and Mobile Computing, IJCSMC, Vol. 2, Issue. 4, April 2013, pg.7 18
- [2] **Abhayadap anawaranita**, garjepaurnima thakareprajakta gundawarpriyanka kulkarni , “Rule Based English T o Marathi Translation Of Assertive Sentence” , International Journal of Scientific & Engineering Research, Volume 4, Issue 5, May 2013
- [3] **M.L Dhore** S.K Dixit, Ruchi M Dhore , “I ssues in Hindi to English and Marathi to English Machine Transliteration of Named Entities” , nternational Journal of Computer Applications (0975 – 8887) Volume 5 No.14, August 2012 .
- [4] **Rajat Kumar Mohanty**, Pushpak Bhattacharyya “Synset Based Multilingual Dictionary: Insights, Applications and Challenges”, [online] <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.409.831&rep=rep1&type=pdf>
- [5] **Nikhil Welankar** , Anirudha Joshi , Kirti Kanitkar , “Principles for Simplifying Translation of Marathi Terms in Mobile Phones” [online] [http://www.bcs.org/upload/pdf/ewic\\_ihci10\\_paper2.pdf](http://www.bcs.org/upload/pdf/ewic_ihci10_paper2.pdf).
- [6] **Jyoti Singh**, Nisheeth Joshi, Iti Mathur , “Part of Speech Tagging for marathi text using Trigram Method” International Journal of Advance Information Technology (IJAIT) Vol. 3, No.2, April 2013.
- [7] **Mangala Madankar** Dr.M.B.Chandak Nekita Chavhan , “nformation Retrieval System and Machine Translation: A Review” International Conference on Information Security and Privacy (ICISP2015), 11-12 December 2015, Nagpur, INDIA.
- [8] Sreelekha .S, Raj Dabre, Pushpak Bhattacharyya, “Comparison of SMT and RBMT the Requirement of Hybridization for Marathi Hindi MT.
- [9] **Pramod Salunkhe**, Mrunal Bewoor, Suhas Patil, Shashank Joshi, Aniket Kadam, “Summarization

and Hybrid Machine Translation System for English to Marathi: A Research Effort in Information Retrieval System(H-Machine Translation)” ,Discovery the International Daily journal ISSN 2278–5469 EISSN 227–545.

- [10] **Pramod Salunkhe** , Mrunal Bewoor , Dr.Suhas Patil Research Work on English to Marathi Hybrid Translation System “A Research Work on English to Marathi Hybrid Translation System”, International Journal of Computer Science and Information Technologies, Vol. 6 (3) 2015, 2557-2560.